Summary and response to the 2013 CIE review of AFSC rockfish

In general, the CIE provided a constructive review and stated that the assessment methods applied were generally “scientifically sound” and of a “high standard”. However, the reviewers highlighted a number of areas which could be improved, ranging from derivation of survey biomass estimates to account for aggregated spatial distributions, estimation of key model parameters such as natural mortality and maturity, the functional form and estimation of selectivity, weighting of data (including reconstructed catch data), and greater consistency in matching the spatial scale of management units with the estimated spatial scale of the stock dynamics. The limited time during the review did not provide an opportunity to closely examine these issues, although one reviewer did comment that the overall assessment results may be expected to be relatively robust under current stock-wide exploitation levels. Improvements could also be made in the presentation of diagnostics and sensitivity analyses. Most of the data-limited criticisms were focused on whether trawl survey biomass estimates sufficiently accounted for aggregated spatial distributions to be considered useful, and offered several suggestions for improvement. For data-limited stocks, using simple time-series approaches to “smooth” a series of survey biomass estimates was recommended, as well as novel approaches such as hierarchical age-structured modeling and more quantitative approaches to evaluating the risks from fishing impacts. Due to the volume of material reviewed, the reviewers strongly recommended that future CIEs would benefit from focusing the review by, for example, geographic region (i.e., GOA or BSAI), model structure (i.e. or age-structured versus non-age structured), or by species.

The following tables show the main recommendations/critiques organized by terms of reference. We respond generally to each major area of the terms of reference below.

Terms of reference

a) Evaluation of data used in the assessments, specifically trawl and longline survey abundance estimates, and recommendations for processing data before use as assessment inputs.

Reviewer	Recommendation
Dichmont, Klaer, Kupschus	Longline data are useful for Tier 3 stocks in the GOA.
Dichmont, Klaer, Kupschus	Tag small shortspine thornyheads
Dichmont, Klaer	Use zero-inflated models for survey data
Dichmont, Klaer	Conduct further work on aging fish (also to get maturity and growth)
Klaer, Kupschus	Investigate sensitivity to errors in catch data

This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the National Marine Fisheries Service and should not be construed to represent any agency determination or policy.

Kupschus	Instead of a zero-inflated approach, consider Poisson errors distribution with GAMs or GLMs.
Kupschus	Models that are linked to habitat data could improve precision but not bias
Dichmont	Add additional surveys in untrawlable grounds
Dichmont	Examine the effect of the change to 15 minutes tows in the trawl survey
Dichmont	Investigate longline survey gear saturation
Klaer	Provide summaries of catch by fishery by region
Klaer	Inclusion of YOY survey indices with better species ID

Survey related recommendations

We agree with the reviewers concerns and recommendations regarding the trawl survey data, particularly for rockfish. Modeling the survey biomass estimates with GAMs or GLMs is a step that should be explored. The recommendation to utilize habitat as a covariate is also good and could greatly improve rockfish biomass estimation. This information could be modeled either within the GAM/GLM approach or other approaches that use statistical distributions that reflect both habitat and species biomass. Ideally, this work should be done in cooperation with the RACE division to leverage their knowledge of the survey data and help ensure the development of a consistent methodology. We hope continued research on the proportion of untrawlable grounds and species densities within those grounds will assist in developing catchability estimates for both biomass based and age-structured assessment models. The issue of the change to 15 minute tows has been explored in the past, but could be revisited when a GAM/GLM modeling approach is initiated. The development of YOY survey indices would be beneficial, but unlikely given the difficulty of adding additional survey time. Environmental indices related to YOY abundance being developed through integrated ecosystem research programs such as the GOA Project may be a more feasible solution to improve estimation of recruitment.

The review also suggested that the authors investigate gear saturation on the longline survey. Some analyses of this were conducted for Mike Sigler's Ph.D. thesis and showed it was negligible for sablefish. However, Rodgveller et al. (2008) presented results that suggest potential for hook competition affecting catch rates for rockfish species and Rodgveller et al. (2011) estimated the relative catchability for the rougheye and blackspotted rockfish complex. We hope to conduct further analyses to assess the extent of possible non-linearities between longline survey indices and stock abundance.

Maturity

The reviewers recommended obtaining more maturity-at-age data to elevate some Tier 5 stocks to Tier 4. We agree that this should be a priority and the maturity work that is ongoing at the Kodiak RACE laboratory has greatly improved our maturity-at-age estimations for GOA rockfish species (e.g., Conrath and Knoth 2013). Solutions for the logistically more challenging problem of obtaining maturity samples in the Aleutian Islands will need to be developed.

Thornyhead tagging

The reviewers suggested that we make an effort to tag smaller shortspine thornyheads, possibly on the trawl survey. While tagging of small thornyheads could be potentially increased by using the trawl survey, their survival would likely be lower than longline caught thornyheads. Additionally, we examined the length composition of tag releases from the longline survey and noted that relatively small (<13 cm) thornyheads are being tagged.

b) Evaluation of analytical methods used in assessments, particularly in regard to selectivity, selection of age and length bin structures, data weighting assumptions, and assumptions and modeling of trawl and longline catchability.

Reviewer	Recommendation
Dichmont, Klaer, Kupschus	For BSAI POP, evaluate dome-shaped selectivity – this would improve consistency with GOA POP (Klaer), and more work should be done to what would cause this (Kupschus).
Dichmont, Kupschus	Develop consistency in age-structured models between regions, or document differences
Dichmont, Klaer	Conduct sensitivity tests as a standard practice
Dichmont, Kupschus	For BSAI POP, fix the residual pattern in the survey composition plus group
Dichmont, Klaer	For BSAI POP, consider estimating growth within the model
Dichmont, Kupschus	For BSAI POP, investigate why the estimate of M is higher than the prior distribution
Dichmont, Klaer	For BSAI POP, fix the slope of the fishery selectivity curve, or use a spline
Dichmont, Klaer	For GOA POP, consider estimating growth within the model
Klaer	Consider and fully justify the exclusion of assessments that assume a stock-recruitment relationship for age-structured assessments.
Kupschus	Weighting of likelihood components a) is not explained/justified; b) differs between models; and c) is very restrictive for catch
Kupschus	The effective sample size weighting in the likelihood needs to be done by sample and not across the sum of the survey
Kupschus	The aging error matrix needs to deal appropriately with the fact that the aging errors in the plus group diminish as the age of the fish moves up into the plus group.
Kupschus	For BSAI POP, little evidence of large cohorts in the 1960s that have been fished hard
Kupschus	For BSAI POP, catch curve estimates of Z are close to assumed M, which implies that the biomass is assumed to be large
Kupschus	Age/size specific M would might explain the high levels of survey plus group
Dichmont	For GOA POP, conduct sensitivity tests for setting the fishery selectivity curve in the most recent block to logistic-gamma

Dichmont	For GOA POP, investigate a non-random pattern in the survey age comp residuals
Klaer	For GOA POP, evaluate sensitivity to alternate break years for fishery selectivity curve
Dichmont	For GOA POP, investigate the appropriate value for M
Kupschus	For GOA POP, it is less clear that that there has been an ontogenetic movement offshore consistent with the shift inshore of the fleet.
Dichmont	For GOA RE/BS complex, evaluate older and larger plus groups
Klaer	Use bridging analysis to show the effect of each incremental change in the assessment
Klaer	Show residual plots and retrospective plots in the assessments
Klaer	Develop SS model to explore different model options (i.e. use of a stock-recruitment relationship)

Consistency between rockfish age-structured models

The population models used in the GOA and BSAI are very similar integrated statistical catch-at-age models. Within each area, a consistent modeling code is used for the various age-structured assessments. The differences in the model code between regions largely reflect choices regarding weighting of likelihood components, whether some parameters should be fixed or estimated (i.e., standard deviation of recruitment deviations), and flexibility in modeling options (i.e., allowing choices in establishing the initial numbers at age, or modeling stock-recruitment curves) rather than fundamentally different modeling frameworks. The similarity of models between regions results from both models originating from the AMAK model. The rockfish modeling workshop held a decade ago (Courtney et al. 2007) describe the general model template.

The weighting of likelihood components across rockfish models reflects both actual uncertainty in the data and legacy weighting from prior authors. One reviewer suggested that the precision assumed for catch is too high. In some cases, authors have accounted for this to some extent by specifying a lower precision for historical estimates of catch. For example, in GOA northern rockfish, the catch is given a lower precision when it is reconstructed as a ratio of POP catch. While we do not have the data to do a better catch reconstruction, we will present sensitivities to the magnitude of historical catch. We intend to further examine data weightings, particularly in terms of effective sample size, and show more diagnostics and sensitivities related to these weightings.

One reviewer suggested that rockfish models could be conducted in Stock Synthesis. In contrast, another reviewer noted that a danger in using pre-packaged models in which the user has less familiarity with the underlying model operations is that errors could be introduced in the input files and/or the model implementation could be different from what is intended. This could result in model interpretation errors. We prefer the flexibility and thorough knowledge of the methodology of our customized models, but agree that attempting to replicate them in Stock Synthesis would be a good validation, and allow for easy sensitivity implementations. We will explore using Stock Synthesis for this purpose in the future.

Ageing error matrix and plus group problems

The reviewers noted a poor fit to the survey age composition plus group for BSAI POP. One reviewer took a more thorough examination of this issue for BSAI POP (and noted that similar issues appeared to exist for other stocks as well but there was insufficient time to examine those cases in detail). Multiple suggestions to improve model fit for the BSAI POP plus group were made, including examining the reliability of the early catch data and recruitment estimates and considering different functional curves for selectivity. All reviewers suggested that dome-shaped fishery selectivity may address the issue for BSAI POP, so this will be pursued in future assessments. More generally, the reviewers also suggested that exploring different values of M, time-specific M, or different selectivities could resolve the residual patterns. It was also recommended that if time-varying selectivity is used that it should be done with as few parameters as possible, such as (for the current BSAI POP model) fixing the slope and allowing age-at-50% selection to vary with a spline. Reviewers supported the GOA POP dome-shaped selectivity justification, but suggested that sensitivity to the placement of blocks and the steep drop in the recent selectivity should be investigated. These issues will be explored in upcoming full assessments.

One issue that the authors need to explore for all the age-structured models is to ensure that the ageing error matrix accounts for a reduced aging error as fish age within the plus group. This occurs because once a fish is well past the accumulator age class, there should be a very small chance that the age assigned by the age reader would not be in the plus group. The BSAI rockfish authors have updated the aging error matrices to address this issue, and the GOA authors will explore this in their upcoming assessments.

Growth and recruitment

Reviewers suggested estimating growth internally in the model. We will explore this possibility, but at the minimum, some of the rockfish models need to update growth and see if there have been temporal changes in growth rate. Slow-growing rockfish typically do not exhibit major interannual fluctuations, but estimating internally should be explored. A drawback of internal estimation is that the estimated growth parameters may be influenced by other modeled parameters (i.e. selectivity, catchability, natural mortality) and potentially lead to inaccurate estimates. On the other hand, estimating growth outside the model from data obtained in trawl surveys typically does not account for the fact that the observed data are filtered through the survey selectivity curve. The decision of whether to estimate growth inside or outside the model would be a trade-off between these issues.

One reviewer noted that the decision not to use a stock-recruitment relationship should be better justified. We do explain in the POP assessments that we do not consider a stock-recruit relationship to be a reasonable assumption, because some of the largest recruitments have come at the lowest stock sizes. This has led us to the assumption that recruitment is more of a function of environmental variability than stock-size (given that stock size does not become extremely low).

Diagnostics, sensitivity, and bridging analysis

The reviewers were in agreement that the rockfish models were lacking in model diagnostics, sensitivity analyses and bridging analyses (the incremental change in the model results with new modeling assumptions and/or new data). Because of the growing size of the stock assessments, we have not been including diagnostics (e.g., residual plots) in the documents unless they were used to highlight an issue. We also do not have any formal sensitivity runs in each document. Bridging analyses that reflect new modeling assumptions are typically included in assessments, whereas a bridging analysis that reflects updated data may be excluded. However, during Plan Team meetings both types of bridging analyses are often shown, particularly when updated data results in substantial changes in estimated model parameters.

Our direction forward to improve in this area is to include sensitivity runs where major parameters (e.g., natural mortality) and data (e.g., catch) are jittered +/- XX% to show whether the models are robust in estimating management quantities such as ABC. A plot or table of these runs will be included in the primary SAFE. For diagnostic plots, and perhaps plots such as length composition fits, we propose to use a web-link to have these available online, but not included in the primary SAFE document. One way to improve our bridging analyses is to include a table that shows incremental model changes over the history of the assessment.

c) Evaluation, findings, and recommendations on the analytic approach used for "data---poor" rockfish stocks and complexes, including the use of an age---structured model for a two--- species complex, and application of state---space production models to stocks and stock complexes.

Reviewer	Recommendation
Dichmont, Klaer, Kupschus	For Tier 4 and 5 species, use Kalman filters or random walk models
Dichmont, Klaer, Kupschus	For low-abundance stocks, apply data-poor risk assessment (PSA or Zhou and Griffiths (2008))
Klaer, Kupschus	Develop MSEs for rockfish
Dichmont	Use zero-inflated methods to produce a survey biomass index
Dichmont	For BS/RE complex, conduct sensitivities for catch data, setting of plus group, and selectivity options
Kupschus	Methods for distinguishing between no survey and no positive catches needs to be developed.
Dichmont	Investigate hierarchical models for data-poor stocks

For data-poor stocks, trawl survey data take on additional importance because they are typically the only information available on the relative stock status. Thus, in these cases it may be especially important to develop survey estimates that address the uncertainty associated with spatially aggregated stocks and overdispersed statistical distributions (i.e., high variance to mean ratios). It may be possible to fit surplus production models to rockfish stocks, which would account for how harvest has affected the stock abundance and thus could provide information on stock productivity. However, the cases where this has been applied (e.g., BSAI blackspotted/rougheye and shortraker rockfish) have not yielded informative parameter estimates on stock productivity and have essentially served as a smoother of the survey data. The reviewers recommended that simple random walk time-series models, potentially implemented through Kalman Filters or random effects models, would be preferred over the practice of taking either a weighted or unweighted average of recent survey biomass estimates. We generally agree, and suggest that surplus production models be at least considered before resorting to time-series models with no underlying population dynamics.

One reviewer noted that the time series models need some procedure for addressing cases in which a survey was conducted but no positive catches occurred, as the variance for this estimate of zero biomass is not defined. The Plan Team survey averaging work group is currently working on this and other similar issues. One potential approach is to use a constant value for survey observation error rather than the time-varying estimates of variability associated with sampling in each survey. Temporal differences in estimated variability of survey biomass estimates that result from changes in availability and/or catchability would be informative to our models, but temporal changes of survey biomass estimates could also represent sampling variability. In general, the work group on survey averaging has considered a variety of methods and will provide a progress report and recommendations at the September 2013 Plan Team meeting.

The application of hierarchical models to data-poor stocks may avoid the situation of assessing stock size and productivity of stocks for which there is little information. In this approach, age-structured models simultaneously fit a complex of stocks that range from data-rich to data-poor, and information from the data-rich stocks is applied to data-poor stocks. Punt et al. (2011) applied this approach to stocks in Australia's scalefish and shark fishery, and assumed commonalities between stocks with respect to fishing selectivity, recruitment deviations, and trends in exploitation rates. While this approach could provide some improvement over current practices, differences between stocks (i.e., from fine-scale targeting that could result in differences in selectivity and exploitation history between stocks) would be expected to degrade performance. We agree that this is an approach that should be investigated further, logically beginning with a management strategy evaluation (MSE). More generally, we agree that MSEs should be developed for Alaska rockfish and can provide a useful tool for both data-rich and data-poor stocks.

Qualitative risk assessments such as Productivity-Susceptibility Analysis (Ormseth and Spencer 2011) have been applied to Alaska stocks, and can provide relative rankings of the vulnerability of fishing within a group of stocks but not necessarily provide estimates of vulnerability in absolute units. One reviewer recommended an approach that developed more quantitative estimates of fishing impacts based on spatial distributions of the stock and fishing effort (Zhou and Griffths 2008). The model depends upon fishery catch rates (i.e., the probability of a fish within a trawl path being captured), which are not likely to be available for non-target rockfish stocks. Nonetheless, we agree that this approach should be pursued and could potentially improve upon the previous PSA analysis conducted for Alaska rockfish.

d) Evaluation, findings, and recommendations on the adequacy of current levels of spatial management, including apportionment strategy.

Reviewer	Recommendation
Dichmont, Klaer, Kupschus	There is evidence of fine-scale structure, and management units should be reasonably small
Dichmont, Kupschus	In genetic modeling by Spies et al., consider how fishing effort would vary spatially as a function of target species density.

Klaer, Kupschus	Use consistent survey averaging procedures for the apportionment process and obtaining overall biomass estimates.
Dichmont	Develop a consistent approach to apportion OFL and ABC across species and regions, with deviations from this approach occurring for a well-specified reason
Dichmont	Consider a series of approaches for smoothing survey time series (i.e., Kalman Filter, Random Effects, models linked with habitat data).
Dichmont	In genetic modeling by Spies et al., consider higher dispersal rates.

The three reviewers were unanimous in their assessment that there is evidence suggesting fine-scale stock structure for Alaska rockfish (<= 500 km), and that the management units should be reasonably small to reflect this structure. One reviewer noted that although the current definitions of “stocks” in either the BSAI or GOA regions are larger than estimated generational dispersal distances based on genetics, there may be insufficient information to develop separate population models at spatial scales smaller than the current stock boundaries. Another reviewer noted that localized depletion was unlikely to affect genetic diversity, but could lower overall stock productivity. The reviewers supported the approach of allocating and monitoring catch among subareas. This recommendation is consistent with recommendations of the NPFMC stock structure working group.

The reviewers noted several inconsistencies in spatially allocating catch. First, the apportionment of catches to subareas differs between regions, with generally finer scale apportionment used in the GOA than in the BSAI. Secondly, the apportionments are not consistent between species; for example several different systems of subarea apportionment are currently applied to BSAI rockfish species. We agree with the reviewers that the reasons for these differences have not been defined, and that a consistent approach that would only change for well-specified reasons should be adopted.

Apportionments are based on the relative spatial distribution of estimated survey biomass, which for Tier 5 stocks is the same survey data used to obtain the area-wide harvest level. However, an additional inconsistency is that in some cases the method of survey averaging used to obtain the area-wide ABC differs from that used to obtain the subarea apportionment. We generally agree that it is appropriate to have consistency in these survey averaging procedures, and further examination of this issue was one of the primary motivations for creating the Plan Team work group on survey averaging.

The reviewers were supportive of research conducted by Ingrid Spies as part of her Ph.D. dissertation, in which she has examined the management implications of location of stock boundaries for stocks with an isolation by distance pattern in genetic diversity. This research supports the intuitive result that stock areas larger than the spatial scale of dispersal that group heavily fished and lightly fished areas into a single management unit can mask the potential impacts on subarea stock sizes. One suggestion was to conduct model simulations with increased dispersal rates. Secondly, in the current model, the distribution of fishing effort is a function of the species density and distance from port (reflecting transportation costs). However, the effort applied to stocks taken as bycatch are more a function of the target species than the bycatch species, and the simulations could be modified to reflect the distribution of effort for bycatch species. We agree that these are sensible modifications for Spies and her colleagues to consider.

The Plan Teams and the Council have become aware of the need to develop procedures for cases in which there appears to be stock structure that is on scales smaller than current management boundaries. Since the rockfish group contains many of the stocks where genetics have shown fine-scale structure, the rockfish authors have initiated much of the discussion on the topic. The current state of development has been to recognize that the potential benefits of more conservative spatial management measures must be weighed against the relative costs of implementing such new management measures. As the primary provider of information on rockfish status to the Council system, the rockfish authors will continue to monitor for deleterious effects on the population dynamics of their stocks and attempt to characterize risks to stock status at biologically relevant spatial scales. We also support additional research be conducted that considers genetics and spatial structure in management strategy evaluations.

e) Recommendations for further improvements a) Evaluation of data used in the assessments, specifically trawl and longline survey abundance estimates, and recommendations for processing data before use as assessment inputs

All recommendations put forward in TOR e were recommended in TORs a-d and are discussed there.

References

Conrath, C.L. and B. Knoth. 2013. Reproductive biology of Pacific ocean perch in the Gulf of Alaska. Mar. Coast. Fish. 5.1: 21-27.

Courtney, D.L., J.N. Ianelli, D.H. Hanselman, and J. Heifetz. 2007. Extending statistical age-structured assessment approaches to Gulf of Alaska rockfish (*Sebastes* spp.). *In* Heifetz, J., DiCosimo J., Gharrett, A.J., Love, M.S, O'Connell, V.M, and Stanley, R.D. (eds.). Biology, Assessment, and Management of North Pacific Rockfishes. Alaska Sea Grant, University of Alaska Fairbanks. pp 429–449.

Ormseth, O.A. and P.D. Spencer. 2011. Analyzing the vulnerability of Alaska groundfish to overfishing. Fisheries Research 112:127-133.

Punt, A.E., Smith, D.C.and Smith, A.D.M. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES Journal of Marine Science 68: 972–981.

Rodgveller, C. J., C. R. Lunsford, and J. T. Fujioka. 2008. Evidence of hook competition in longline surveys. Fish. Bull. 106: 364-374.

Rodgveller, C.J., M.F. Sigler, D.H. Hanselman, and D.H. Ito. 2011. Sampling Efficiency of Longlines for Shortraker and Rougheye Rockfish Using Observations from a Manned Submersible, Marine and Coastal Fisheries, 3:1, 1-9.

Zhou, S.J. and S.P. Griffiths, S.P. 2008. Sustainability Assessment for Fishing Effects (SAFE): A new quantitative ecological risk assessment and its application to elasmobranch bycatch.in an Australian trawl fishery. Fisheries Research 91:56–68.